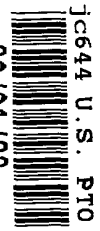
04-03-00

A



Express Mail No. EL5170006315US

Attorney's Docket No. NC17353 **Patent**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Box Patent Application
Assistant Commissioner for Patents
Washington, D.C. 20231

NEW APPLICATION TRANSMITTAL

Transmitted herewith for filing is the patent application of

Inventor(s): Rajeev Koodli

WARNING: Patent must be applied for in the name(s) of all of the actual inventor(s). 37 CFR 1.41(a) and 1 53(b).

For (title): METHOD AND APPARATUS FOR RANDOM PACKET MARKING FOR DIFFERENTIATED SERVICES

CERTIFICATION UNDER 37 CFR 1.10

I hereby certify that this New Application Transmittal and the documents referred to as enclosed therein are being deposited with the United States Postal Service on this date, 3/31/00 , in an envelope as "Express Mail Post Office to Addressee" Mailing Label Number EL517006315US, addressed to the: Assistant Commissioner for Patents, Washington, D.C. 20231.

Beverly Sparkman
(type or print name of person mailing paper)

Beverly Sparkman
Signature of person mailing paper

NOTE: Each paper or fee referred to as enclosed herein has the number of the "Express Mail" mailing label placed thereon prior to mailing. 37 CFR 1.10(b).

WARNING: Certificate of mailing (first class) or facsimile transmission procedures of 37 CFR 1.8 cannot be used to obtain a date of mailing or transmission for this correspondence.

1. Type of Application

This new application is for a(n)
(check one applicable item below)

☒ Original (nonprovisional)

☐ Design
☐ Plant

WARNING: *Do not use this transmittal for a completion in the U.S of an International Application under 35 U.S C 371(c)(4), unless the International Application is being filed as a divisional, continuation or continuation-in-part application*

WARNING. *Do not use this transmittal for the filing of a provisional application.*

NOTE: *If one of the following 3 items apply, then complete and attach ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF A PRIOR U.S. APPLICATION CLAIMED and a NOTIFICATION IN PARENT APPLICATION OF THE FILING OF THIS CONTINUATION APPLICATION.*

☐ Divisional.

☐ Continuation.

☐ Continuation-in-part (C-I-P).

2. Benefit of Prior U.S. Application(s) (35. U.S.C. 119(e), 120, or 121)

NOTE: *If the new application being transmitted is a divisional, continuation or a continuation-in-part of a parent case, or where the parent case is an International Application which designated the U.S., or benefit of a prior provisional application is claimed, then check the following item and complete and attach ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION(S) CLAIMED.*

WARNING: *If an application claims the benefit of the filing date of an earlier filed application under 35 U.S.C. 120, 121 or 365(c), the 20-year term of that application will be based upon the filing date of the earliest U.S. application that the application makes reference to under 35 U.S.C. 120, 121 or 365(c). [35 U.S.C. 154(a)(2) does not take into account, for the determination of the patent term, any application on which priority is claimed under 35 U.S.C 119, 365(a) or 365(b).] For a c-i-p application, applicant should review whether any claim in the patent that will issue is supported by an earlier application and, if not, the applicant should consider canceling the reference to the earlier filed application The term of a patent is not based on a claim-by-claim approach. See Notice of April 14, 1995, 60 Fed. Reg. 20,195, at 20,205.*

WARNING: *When the last day of pendency of a provisional application falls on a Saturday, Sunday, or Federal holiday within the District of Columbia, any nonprovisional application claiming benefit of the provisional application must be filed prior to the Saturday, Sunday, or Federal holiday within the District of Columbia. See 37 C F R. § 1 78(a)(3)*

☒ The new application being transmitted claims the benefit of prior U.S. application(s). Enclosed are ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION(S) CLAIMED.

3. Papers Enclosed That Are Required for Filing Date under 37 C.F.R. 1.53(b) (Regular) or 37 C.F.R. 1.153 (Design) Application

15 Pages of specification

6 Pages of claims

1 Page of Abstract

6 Sheets of drawing

☐ formal

☒ informal

WARNING: ***DO NOT*** *submit original drawings. A high quality copy of the drawings should be supplied when filing a patent application. The drawings that are submitted to the Office must be on strong, white, smooth, and non-shiny paper and meet the standards according to § 1.84. If corrections to the drawings are necessary, they should be made to the original drawing and a high-quality copy of the corrected original drawing then submitted to the Office. Only one copy is required or desired. Comments on proposed new 37 CFR 1.84. Notice of March 9, 1988 (1990 O.G. 57-62).*

NOTE "Identifying indicia, if provided, should include the application number or the title of the invention, inventor's name, docket number (if any), and the name and telephone number of a person to call if the Office is unable to match the drawings to the proper application. This information should be placed on the back of each sheet of drawing a minimum distance of 1.5 cm (5/8 inch) down from the top of the page." 37 C.F.R. 1.84(c).

(complete the following, if applicable)

☐ The enclosed drawing(s) are photograph(s), and there is also attached a "PETITION TO ACCEPT PHOTOGRAPH(S) AS DRAWING(S)." 37 C.F.R. 1.84(b).

4. Additional papers enclosed

☐ Preliminary Amendment

☐ Information Disclosure Statement (37 C.F.R. 1.98)

☐ Form PTO-1449

☐ Citations

☐ Declaration of Biological Deposit

☐ Submission of "Sequence Listing," computer readable copy and/or amendment pertaining thereto for biotechnology invention containing nucleotide and/or amino acid sequence.

☐ Authorization of Attorney(s) to Accept and Follow Instructions from Representative

☐ Special Comments

☐ Other

5. Declaration or oath

☐ Enclosed

Executed by

(check all applicable boxes)

☐ inventor(s).

☐ legal representative of inventor(s). 37 CFR 1.42 or 1.43.

☐ joint inventor or person showing a proprietary interest on behalf of inventor who refused to sign or cannot be reached.

☐ This is the petition required by 37 CFR 1.47 and the statement required by 37 CFR 1.47 is also attached. See item 13 below for fee.

☒ Not Enclosed.

WARNING: Where the filing is a completion in the U S of an International Application, but where a declaration is not available, or where the completion of the U S. application contains subject matter in addition to the International Application, the application may be treated as a continuation or continuation-in-part, as the case may be, utilizing ADDED PAGE FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION CLAIMED.

☐ Application is made by a person authorized under 37 CFR 1.41(c) on behalf of ***all*** the above named inventor(s).

[The declaration or oath, along with the surcharge required by 37 CFR 1.16(e) can be filed subsequently.]

NOTE: It is important that all the correct inventor(s) are named for filing under 37 CFR 1.41(c) and 1.53(b).

☐ Showing that the filing is authorized.
[not required unless called in question. 37 CFR 1.41(d)]

6. Inventorship Statement

WARNING: *If the named inventors are each not the inventors of all the claims an explanation, including the ownership of the various claims at the time the last claimed invention was made, should be submitted.*

The inventorship for all the claims in this application are:

☐ The same.

or

☐ Not the same. An explanation, including the ownership of the various claims at the time the last claimed invention was made,

☐ is submitted.

☐ will be submitted.

7. Language

NOTE: An application including a signed oath or declaration may be filed in a language other than English. A verified English translation of the non-English language application and the processing fee of $130.00 required by 37 CFR 1.17(k) is required to be filed with the application, or within such time as may be set by the Office. 37 CFR 1.52(d).

NOTE: A non-English oath or declaration in the form provided or approved by the PTO need not be translated. 37 CFR 1.69(b).

☒ English

☐ Non-English

☐ The attached translation is a verified translation. 37 CFR 1.52(d).

8. Assignment

☐ An assignment of the invention to Nokia Mobile Phones Limited

☐ is attached. A separate ☐ "COVER SHEET FOR ASSIGNMENT (DOCUMENT) ACCOMPANYING NEW PATENT APPLICATION" or ☐ FORM PTO 1595 is also attached.

☒ will follow.

NOTE: "If an assignment is submitted with a new application, send two separate letters–one for the application and one for the assignment." Notice of May 4, 1990 (1114 O.G. 77-78).

WARNING: A newly executed "CERTIFICATE UNDER 37 CFR 3.73(b)" must be filed when a continuation-in-part application is filed by an assignee. Notice of April 30, 1993, 1150 O.G. 62-64.

9. Certified Copy

Certified copy(ies) of application(s)

Country	Appln. no.	Filed
Country	Appln. no.	Filed
Country	Appln. no.	Filed

from which priority is claimed

☐ is (are) attached.

☐ will follow.

NOTE: The foreign application forming the basis for the claim for priority must be referred to in the oath or declaration. 37 CFR 1.55(a) and 1.63.

NOTE: This item is for any foreign priority for which the application being filed directly relates. If any parent U.S. application or International Application from which this application claims benefit under 35 U.S.C. 120 is itself entitled to priority from a prior foreign application, then complete item 18 on the ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION(S) CLAIMED.

10. Fee Calculation (37 C.F.R. 1.16)

A. ☒ Regular application

CLAIMS AS FILED			
Number Filed	Number Extra	Rate	Basic Fee 37 C.F.R. 1.16(a) $690.00
Total Claims [37 CFR 1.16(c)] 23-20 = 3	3	x $18.00	54.00
Independent Claims [37 CFR 1.16(b)] 3-3 = 0	0	x $78.00	0
Multiple dependent claim(s), if any [37 CFR 1.16(d)]	0	+ $260.00	0

☐ Amendment cancelling extra claims is enclosed.

☐ Amendment deleting multiple-dependencies is enclosed.

☐ Fee for extra claims is not being paid at this time.

NOTE: If the fees for extra claims are not paid on filing they must be paid or the claims cancelled by amendment, prior to the expiration of the time period set for response by the Patent and Trademark Office in any notice of fee deficiency. 37 CFR 1.16(d).

Filing Fee Calculation $744.00

B. ☐ Design application
[$310.00—37 CFR 1.16(f)]

Filing Fee Calculation $

C. ☐ Plant application
[$480.00—37 CFR 1.16(g)]

Filing Fee Calculation $

11. Small Entity Statement(s)

☐ Verified Statement(s) that this is a filing by a small entity under 37 CFR 1.9 and 1.27 is (are) attached.

WARNING: *Status as a small entity in one application or patent does not affect any other application or patent, including applications or patents which are directly or indirectly dependent upon the application or patent in which the status has been established. A nonprovisional application claiming benefit under 35 U.S.C. 119(e), 120, 121 or 365(c) of a prior application may rely on a verified statement filed in the prior application if the nonprovisional application includes a reference to a verified statement in the prior application or includes a copy of the verified statement filed in the prior application if status as a small entity is still proper and desired." 37 C F.R. § 1.28(a).

(complete the following, if applicable)

☐ Status as a small entity was claimed in prior application.

_____/_____, was filed on _____, from which benefit is being claimed for this application under:

35 U.S.C. ☐ 119(e),
☐ 120,
☐ 121,
☐ 365(c),

and which status as a small entity is still proper and desired.

☐ A copy of the verified statement in the prior application is included.

Filing Fee Calculation (50% of A, B or C above)

$_____

NOTE: *Any excess of the full fee paid will be refunded if a verified statement and a refund request are filed within 2 months of the date of timely payment of a full fee The two-month period is not extendible under § 1.136, 37 CFR 1.28(a).*

12. Request for International-Type Search [37 C.F.R. 1.104(d)]

(complete, if applicable)

☐ Please prepare an international-type search report for this application at the time when national examination on the merits takes place.

13. Fee Payment Being Made at This Time

☐ Not Enclosed

☐ No filing fee is to be paid at this time.
(*This and the surcharge required by 37 C.F.R. 1.16(e) can be paid subsequently).*

☒ Enclosed

☒ Filing fee	$744.00
☐ Recording assignment [$40.00; 37 C.F.R. 1.21(h)] (See attached "COVER SHEET FOR ASSIGNMENT ACCOMPANYING NEW APPLICATION").	
☐ Petition fee for filing by other than all the inventors or person on behalf of the inventor where inventor refused to sign or cannot be reached [$130.00; 37 C.F.R. 1.47 and 1.17(h)]	
☐ For processing an application with a specification in a non-English language [$130.00; 37 C.F.R. 1.52(d) and 1.17(k)]	
☐ Processing and retention fee [$130.00; 37 C.F.R. 1.53(d) and 1.21(l)]	
☐ Fee for international-type search report [$40.00; 37 C.F.R. 1.21(e)]	

NOTE: 37 CFR 1.21(l) establishes a fee for processing and retaining any application that is abandoned for failing to complete the application pursuant to 37 CFR 1.53(d) and this, as well as the changes to 37 CFR 1.53 and 1.78, indicates that in order to obtain the benefit of a prior U.S. application, either the basic filing fee must be paid, or the processing and retention fee of § 1.21(l) must be paid, within 1 year from notification under § 53(d).

Total fees enclosed $744.00

14. Method of Payment of Fees

☐ Check in the amount of $_____

☒ Charge Account No. 50-0270 in the amount of $744.00

Two duplicates of this transmittal are attached.

NOTE: *Fees should be itemized in such a manner that it is clear for which purpose the fees are paid. 37 CFR 1.22(b).*

15. Authorization to Charge Additional Fees

WARNING: If no fees are to be paid on filing, the following items should not be completed

WARNING: Accurately count claims, especially multiple dependent claims, to avoid unexpected high charges, if extra claim charges are authorized.

☒ The Commissioner is hereby authorized to charge the following additional fees by this paper and during the entire pendency of this application to Account No. 50-0270.

- ☒ 37 C.F.R. 1.16(a), (f) or (g) (filing fees)
- ☒ 37 C.F.R. 1.16(b), (c) and (d) (presentation of extra claims)

NOTE: *Because additional fee for excess or multiple dependent claims not paid on filing or on later presentation must only be paid or these claims cancelled by amendment prior to the expiration of the time period set for response by the PTO in any notice of fee deficiency [37 CFR 1.16(d)], it might be best not to authorize the PTO to charge additional claim fees, except possibly when dealing with amendments after final action.*

- ☒ 37 C.F.R. 1.16(e) (surcharge for filing the basic filing fee and/or declaration on a date later than the filing date of the application)
- ☒ 37 C.F.R. 1.17 (application processing fees)

WARNING: While 37 CFR 1.17(a), (b), (c) and (d) deal with extensions of time under § 1 136(a), this authorization should be made only with the knowledge that "Submission of the appropriate extension fee under 37 C.F.R. 1 136(a) is to no avail unless a request or petition for extension is filed." (Emphasis added). Notice of November 5, 1985 (1060 O.G 27).

- ☐ 37 C.F.R. 1.18 (issue fee at or before mailing of Notice of Allowance, pursuant to 37 C.F.R. 1.311(b))

NOTE: *Where an authorization to charge the issue fee to a deposit account has been filed before the mailing of a Notice of Allowance, the issue fee will be automatically charged to the deposit account at the time of mailing the notice of allowance. 37 CFR 1.311(b).*

NOTE: *37 CFR 1.28(b) requires "Notification of any change in loss of entitlement to small entity status must be filed in the application...prior to paying, or at the time of paying,...issue fee." From the wording of 37 CFR 1.28(b): (a) notification of change of status must be made even if the fee is paid as "other than a small entity" and (b) no notification is required if the change is to another small entity.*

16. Instructions as to Overpayment

- ☒ Credit Account No. 50-0270.
- ☐ Refund

Robert C. Rolnik

Signature of Attorney

Reg. No. 37,995

Robert C. Rolnik
(type or print name of attorney)

Tel. No. (972) 894-5931

Nokia Inc.
6000 Connection Drive 1-4-755
(P.O. Address)
Irving, TX 75039

☒ **Incorporation by reference of added pages**

[check the following item if the application in this transmittal claims the benefit of prior U.S. application(s) (including an International Application entering the U.S. stage as a continuation, divisional or C-I-P application) and complete and attach the ADDED PAGES FOR NEW APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION(S) CLAIMED.]

☒ Plus Added Pages for New Application Transmittal Where Benefit of Prior U.S. Application(s) Claimed

Number of pages added 1

☐ Plus Added Pages for Paper Referred to in Item 4 Above

Number of pages added _____

☐ Plus "Assignment Cover Letter Accompanying New Application"

Number of pages added _____

☐ **Statement Where No Further Pages Added**

(if no further pages form a part of this Transmittal, then end this transmittal with this page and check the following item)

☐ This transmittal ends with this page.

EL517006315US

Attorney Docket No. NC17353 ***PATENT***

ADDED PAGES FOR APPLICATION TRANSMITTAL WHERE BENEFIT OF PRIOR U.S. APPLICATION(S) CLAIMED

35 U.S.C. 119(e)

This application claims the benefit of U.S. Provisional Application(s) No(s).:

APPLICATION NO(S).	**FILING DATE**
60/159,522	10/15/99

Attorney Docket No. NC17353

Patent Application Papers of: Rajeev Koodli

METHOD AND APPARATUS FOR MARKING DATA PACKETS IN A DIFFERENTIATED SERVICES NETWORK

FIELD OF THE INVENTION

This invention relates generally to routing packets through a communication network of network entities, and more specifically, to a manner of marking packets from several data flows so as to achieve greater fairness in the further propagating of such packets through the communication network.

BACKGROUND OF THE INVENTION

In the Internet Engineering Task Force's (IETF) Differentiated Services architecture (S. Blake et al., "An Architecture for Differentiated Services", IETF RFC 2475) for providing IP Quality of Service (QoS), each IP packet carries a Differentiated Services Code-Point (DSCP) for the Differentiated Services (DS) field. DSCP is an index into a list of Per-Hop Behaviors (PHBs) that a packet may be entitled to at each DS-compliant node or router. A PHB may include a probability or preference to drop a packet of a certain class. By obtaining similar PHB at each network node using the DSCP as a tag or marker, an IP flow can realize end-to-end QoS.

In the IETF model, a source sends packets to a network, which may have wireless links. A first-hop router, also known as ISP router, places the appropriate DSCP in the DS field of each packet. If the source sends traffic, comprised of data packets in a flow, according to an agreed contract or policy, the packets are considered "in-profile" and marked with an appropriate DSCP. For example, those packets that are received according to the agreed contract, may be marked with a selected priority level from among several

priority levels. Each priority level may have at least one DSCP associated with the priority level such that no DSCP is assigned more than a single priority level. Each of the priority levels denote a drop precedence, or color of packet. A high priority level is less likely to be dropped than a lower priority level. A subordinate priority level is any priority level that is not the highest priority level.

In the Three Color Marker (TCM) model (see J. Heinanen and R. Guerin, "A Two Rate Three Color Marker", Internet Draft, May 1999), a packet marked with a priority level of green is least likely to be dropped at a router, as compared to packets marked with priority levels of yellow or red. In the TCM model, the highest priority level is green. Red and yellow are subordinate priority levels. Thus it is advantageous in a network of network entities or routers that use multiple priority marked packets, e.g. TCM, for a user to have as many data packets in a flow marked green as possible, at least from the point of view of that user.

Unfortunately, since the resources at each router, and of the network generally, are limited, what is good for a single user, may have a negative impact on other users. If a single user monopolized the entire buffer queue at a router, that would leave the router unavailable to other users who intend to use the router. Marking a preponderance of data packets of a first user green, while marking a minority of data packets of a second user as green, would have a similar, but less pronounced effect. In that situation, a smaller percentage of the first user's packets would be dropped, (since green is low drop probability) as compared to the packets of the second user, which has a greater percentage of packets marked with the inferior priority levels of red and yellow.

Fairness is a measure of proportionally marking packets (with different colors corresponding to different packet-drop precedences) originating from a user of a customer consuming some bandwidth X, wherein the proportion of packets marked for the user of the customer is as close as possible to the

proportion of packets marked of a second user of the customer, even though the second user may currently consume a bandwidth different than X. Necessarily, the proportion of packets marked, is a measure of the number of packets from a user marked over a period of time, compared to the overall set of packets of that user being considered for marking over the same period of time.

TCM tends to permit high priority marking proportions to vary considerably between users of a common customer, when compared at the same time. Such a disparity can lead to dissatisfaction of users, not entirely unlike that experienced by motorists who encounter a traffic jam.

SUMMARY OF THE INVENTION

A method of transmitting data on a network achieves greater fairness between packet flows from different sources. The proportion of packets to be marked a priority level, may be determined, in part, by a rate threshold.

An ability to establish a credit for good behavior, such as underutilized capacity, is achieved. The credit permits occasional bursts of packets above a threshold while the credit continues to satisfy a criterion. Such credits may be shared among several packet flows.

A further object of the invention is a means to adjust a probability of selecting a priority level on the basis of weighting multiple factors. This may enable a network operator to adjust between a preference for short duration adherence to a rule, and a preference for long duration adherence to a rule.

According to an embodiment of the invention, packets from a source reach a router. The router determines a sending rate estimate. The packet is then marked with a priority level based on the sending rate estimate.

The embodiment of the invention, may be operated in a diffserv or other network environment where one or more users, bargains with a ISP for carriage of packets over the network. The bargain struck results in setting parameters or rate thresholds to govern the rates of packets originating from a customer domain, wherein the one or more users may transmit packets from within IP addresses of the customer domain. Each user may have one or more IP flows associated with the user. An IP flow may be characterized by the source IP address, the destination IP address, the port numbers and the protocol id. Similar parameters may be used if the flow is an IPv6 flow. The sum of the packets of all users of a customer is known as the aggregate flow, or just 'aggregate'. An example of a customer would be Nokia Corporation. Another example would be a campus at a university, wherein the users may include faculty, staff and students. In the end, the customer may be regarded as a collective that has bargained for certain packet transmission qualities, and in particular, for rates of transmission and packet-drop probability. A user in this context is anyone who is apparently authorized to operate the equipment within the customer's domain that generates a flow.

A packet marker embodiment, known as a Random Packet Marker (RPM) marks packets on a flow-aggregate or aggregate basis rather than on per-flow basis. The marking is done as a function of the packet sending rate of the aggregate with respect to at least one rate threshold, e.g. a Committed Information Rate (CIR), established by prior agreement between the customer and the ISP. When the sending rate is at or below the CIR, all packets are sent as green; when it diverges from CIR, the probability of a packet being marked as green decreases, while the probability of being marked as yellow or red or other lower priority increases. A super rate threshold is one that is higher than at least one other rate threshold.

In comparison to TCM, testing shows that RPM provides a greater level of fairness across multiple flows, wherein fairness may be measured as

the standard deviation of the proportion of packets marked green among the multiple flows.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block diagram of data flows traveling over links to a router according to an embodiment of the invention;

Fig. 2 is proportionality diagram showing the relative proportions of color markings of packets under varying rate multiples of the Committed Information Rate (CIR) according to an embodiment of the invention;

Fig. 3 is proportionality diagram showing the relative proportions of color markings of packets under varying rate multiples of the Committed Information Rate (CIR) according to another embodiment of the invention;

Fig. 4 is a simulation model for generating simulated results of an embodiment of the invention;

Fig. 5 is a comparison of a window length to a packet duration of a Time Sliding Window (TSW);

Fig. 6 is a proportionality diagram of a simulation of packet marking for a single TCP source;

Fig. 7 is a proportionality diagram of a simulation of a packet marking for six TCP sources using an embodiment of the invention; and

Fig. 8 is a proportionality diagram of simulated result of a packet marking for six TCP sources using a prior art method and apparatus.

Fig. 9 is a proportionality diagram of a simulated result of a packet marking for a single TCP source using another embodiment of the invention.

DETAILED DESCRIPTION OF THE INVENTION

An embodiment of the invention is shown in Figure 1. An ISP router **101**, operated by an internet service provider (ISP), may have multiple input interfaces belonging to different customers, e.g. input interface i0 **103** and input interface i1 **105**. For example, an AF class denoted by flow x **111**, and flow y **113** may be specified to handle a first rate threshold, e.g. a Committed Information Rate (CIR), of 0.3 Mbps, and a second rate threshold, e.g. Peak Information Rate (PIR), of 0.4 Mbps. In this embodiment, PIR is a super rate threshold. Other higher rate thresholds could also be implemented. EF class traffic is denoted by flow z **115**. The ISP router **101** employing differentiated services has to meter, using methods known in the art, the incoming traffic on the input interface i0 **103**, and mark the packets appropriately based on traffic compliance. The ISP router **101** need not consider whether each flow belonging to a customer is individually traffic compliant; instead, it may consider whether the entire customer traffic, e.g. flow x **111** and flow z **115**, is compliant to the CIR of the customer and PIR of the customer or not. In such case, it is important to be able to provide per-aggregate marking algorithm in a router so that downstream routers are able to discard packets in accordance with the selected priority level set by a packet marker embodiment of the invention.

The packet marker of an embodiment relies on the availability of sending rate information, typically provided by a meter or metering tool. Such metering tools may include a time sliding window, which provides an estimate of sending rates, for example the aggregate sending rate of a customer, sometimes called s.

An embodiment of the invention, called random packet marking (RPM) may operate according to the following rules, wherein green, yellow and red

denote unique DSCP bit patterns available to be set in a data packet, and the probabilities are for individual packets, based on rates determined by, e.g. Time Sliding Window (TSW):

Green marking probability equals:

1 if s <= CIR, and

CIR / s otherwise.

Yellow marking probability equals:

0 if s <= CIR;

1 – (CIR / s) if CIR < s <= PIR; and

(PIR - CIR)/s otherwise.

Red marking probability equals:

0 if s <= PIR; and

(s - PIR)/s otherwise.

Fig. 2 is proportionality diagram showing the relative proportions of color markings of packets. The diagram is a proportionality diagram because proportions are measured on the vertical axis. The diagram shows the likelihood of marking a packet for each of the several priority levels according to the previously mentioned embodiment of the invention. The diagram also shows the likelihood of dropping packets **201** by a diffserv router downstream from the marker, so that although the proportions shown for **203**, **205**, and **207** are local to the node, **201** is the proportion collectively of packets dropped at one of possibly several nodes. In this example the PIR is set to be twice the CIR. The likelihood of dropping packets is known as the overall drop precedence **201**. It may be calculated by:

Overall Drop Precedence = low * $P_G(x)$ + medium * $P_Y(x)$ + high * $P_R(x)$,

where low, medium and high represent numeric probabilities of packet dropping associated with each of the three selected priority levels, and $P_G(x)$ **203**, $P_Y(x)$ **205** and $P_R(x)$, **207** are the probabilities of marking green, yellow and red, respectively.

Fig. 3 is a proportionality diagram that shows operation of another embodiment of the invention. The embodiment may permit greater flexibility, by reducing the overall drop precedence of the previously described embodiment. The second embodiment modifies marking rules to bias, or improve the probability of marking green, while there is a sufficiently accumulated burst-size, b, built up for the customer (aggregate) and a first burst criterion **301** is satisfied, e.g. that s> CIR. The first burst criterion may also include a requirement that burst-size, b, is greater than a minimum burst, e.g. that b > 0. By increasing the proportion of packets marked green, the best priority level, the overall drop rate at downstream routers is reduced. The burst-size, b, may be regarded as a credit to a customer for operating at rates below CIR. And like the first embodiment, the burst-size is accumulated for each customer so that flows of the customer are treated with high fairness. The practical effect this has in relation to Fig. 3, is that it increases the probability, to high, of marking a packet green, $P_G(x)$, to a burst-mode probability **303b**, or high, while the burst credit, or burst-size remains above a threshold, e.g. 0. Naturally, if the probability of marking green, $P_G(x)$, is increased during this time, the probability of marking the packet a lower priority level **305b**, e.g. yellow, $P_Y(x)$, must be reduced accordingly so that the sum of the probabilities is 1. An intermediate burst-mode probability $P_G(x)$ **303c** of marking green may be preferred if the ISP does not want to fix the burst-mode probability to 1 for the highest priority level or green. Similarly, the intermediate burst-mode probability for yellow $P_Y(x)$ **305c** may also be used.

If a sending rate satisfies a second burst criterion **302**, the probability of marking green may decline, as sending rate increases, under a secondary function **313b**. The secondary function **313b** according to the second embodiment may be in proportion to the inverse of the sending rate. The secondary function may be selected by a second burst criterion, such as, upon determining that s is as large as or much larger than PIR and burst-size is above a threshold.

Alternatively, the random element may be removed from the decision to mark green, and the packets may be marked green at a rate not exceeding the PIR or some multiple thereof. The selection of the remaining priority levels, e.g. red and yellow, for marking could be based on the probabilities provided under the operation of the first embodiment.

The burst credit may be established at a level, and may be restricted to a range, wherein no bursts are accumulated above the range, and no bursts are deducted below the range. The burst-size may be incremented in a fixed chunk, or variable chunk, and the burst size may be decremented in a fixed chunk or variable chunks. A suitable chunk setting could be the difference between the sending rate (expressed in bits per second) and CIR (expressed in bits per second) multiplied by inter-pkt-time. Inter-pkt-time may be the time measured between packets received. Inter-pkt-time may be a weighted average of times between several packet arrivals. Inter-pkt-time is also known as inter packet spacing. Another suitable chunk setting could be the number of bits in the packet to be marked.

The burst-size may be decremented by a chunk any time one or more of first burst criterion and second burst criterion is satisfied. The burst size may be incremented a chunk any time all burst criteria are not satisfied.

A more specific example of marking a received packet according to the second embodiment follows, using the language of TCM, wherein green is the

highest priority level. All references of burst-size; sending rate; CIR; PIR; inter-pkt-time; pkt-size apply to a particular customer.

If the sending rate of the customer associated with the packet, s < CIR, then

mark the packet green and

burst-size = burst-size + (CIR – sending rate) * inter-pkt-time.

If CIR < s < PIR and burst-size > 0, then

mark the packet green; and

burst-size = burst-size – pkt-size.

If sending-rate > PIR and burst-size > 0, then

mark the packet so that it and any prior packet marked green for the customer do not exceed the peak information rate; and

burst-size = burst-size – pkt-size.

The values for (CIR - sending rate) * inter-pkt-time and pkt-size may each operate as a chunk. In order to mark packets green while not exceeding the PIR, a count of packets marked green by the marker may be maintained for an appropriately sized time period, or duration. A count of too many green packets in that period, would prohibit the current packet from being marked green, even though the burst-size, b, is greater than zero.

A convenient measure of a stream of data, is the instantaneous sending rate, which may be determined by a meter. Instantaneous sending rate may be determined by identifying two packets and dividing the data carried in the first packet by the duration between packet arrivals. The instantaneous sending rate may be in relation to packets fitting a certain

criteria, such as, e.g. marked as green, in which case it is called the green instantaneous sending rate. The instantaneous sending rate may be a helpful gauge to measure performance of a marker, or of the network in general. In a sense, the instantaneous sending rate is a microscopic view of very few packets at some stage in the process of sending, receiving, metering or marking of packets.

It is possible when using some marking algorithms for the instantaneous sending rate for green packets to be higher than the CIR. Generally, this can occur with a minority of the green packets, but nevertheless the inter-pkt-time, or inter-packet spacing, between two packets may be shorter than an inter-packet spacing derived from the CIR, giving the appearance that a customer is getting a free ride beyond what was bargained for. To strictly prohibit green marked packets from having inter-packet spacing this small (i.e. in violation of the inter-packet spacing set by CIR), a marker could remark those packets marked green by the earlier embodiments. However, a packet marker doing this suffers the problem that averaging the data rate for all packets marked green often results a green rate well under the CIR. An example of this is in the situation where packets uniformly arrive every 8 ms, and minimum spacing, of CIR, is 27 ms. By remarking so that no more than 1 in four consecutive packets are green, the green inter-packet spacing is 32 ms, nearly 25% slower than CIR -- clearly the customer is getting shortchanged when the incoming packets arrive just barely under the inter-packet spacing of CIR.

In order to avoid short-changing the customer, and permit longer term average rates to be closer to the sending rate CIR, a probability of keeping the green packets green using a soft inter-packet spacing step according to an embodiment of the invention would be:

$$P'(\text{green}) = \exp\left((x-1) / a \right)$$

where x is the instantaneous green sending rate divided by CIR prior to remarking, and the remarking occurs provided the instantaneous green sending rate is larger than CIR. If instantaneous sending rate of two green packets prior to remarking does not satisfy this, then P'(green) = 1.

Soft inter-packet spacing, as generated by the soft inter-packet function, P'(green), can be helpful when data from a customer just starts up. Starting up the flows from the customer causes the aggregate sending rate (s) to fluctuate because no or few packets have arrived upon which to estimate s. At that time, an ISP may want a blend of the RPM, and the soft inter-packet spacing embodiments. In addition, the ISP may want to provide the customers rates according to soft inter-packet function, during times when traffic on the ISP is operating well below capacity, i.e. during off-peak hours.

The probabilities produced by the two embodiments -- a first probability in the case of the RPM embodiment, and a second probability in the case of soft inter-packet spacing embodiment -- may be blended together to provide a blended probability by weighting the probability for marking a packet according to each of the embodiments according to table 1.

Duration of use	*Embodiment: P_G (RPM)*	*Embodiment: P' (Soft inter-packet spacing)*
At start-up	.01 weighting	.99 weighting
5 seconds after start-up	.50	.50
30 seconds after start-up	.95	.05
Steady state	1.00	0.00

Table 1

The blended probability, P_b, thus becomes:

$$P_b = w * P_G(x) + (1 - w) * P'(green);$$

where w may range between 0 and 1, and operates to weight each probability algorithm. Such a blend may be helpful to initialize the marking using P'(green) and later shift to $P_G(x)$, i.e. setting w=0 at startup, and shifting to a higher value of w as time passes.

Fig. 4 shows a simulation model that is useful to compare the results of an embodiment with TCM. GRED **401** is a generalized RED, (S. Floyd and V. Jacobson, "Random Early Detection Gateways for Congestion Avoidance", IEEE/ACM Transactions on Networking, August 1993) for handling traffic with multiple drop preferences. Packet source 1 or user 1 **403** provides at least one packet flow. Packet source n or user n **405**, which may be one of several packet sources, also provides at least one packet flow. Sink **407** represents a network entity, e.g. a router or client, that receives the packet flows from packet source 1 **403** and packet source n **405**. The metering method used for the simulation is Time Sliding Window (TSW) with Exponential Weighted Moving Average (EWMA). In TSW, the sending rate may be calculated as :

$$curr\_rate = \frac{win\_length * avg\_rate + pkt\_size}{win\_length + pkt\_time}$$

where, in accordance with Fig. 5, win_length **501** is the past time taken into consideration for calculating the current sending rate. The relationship between win_length **501** and pkt_time **503** is shown in Fig. 5. The variable avg_weight, may be initialized at a default value, e.g. 0, and then modified as

time goes by to be: avg_rate = (1-w) * curr_rate + w * avg_rate; where w is some suitable weight so that more recent packets influence the value for avg_rate more than older packets.

Fig. 6 shows a single TCP flow, broken down into its constituent proportional representation of packets marked red, yellow and green. This kind of graph is a proportionality graph and is helpful to show the fluctuations over time of the proportion of the set of packets that are being marked with each color. One way to calculate the proportion for a given color, is to add up the total packets marked the given color over a duration, say long enough for 20 packets to be marked, then dividing by the total packets of the TCP flow for which there has been an opportunity to mark during that duration. If 3 packets are marked green, during an interval where 20 packets for the flow have been received, then the proportion of packets marked green is 0.15, or 15%. Fig. 6 shows a proportionality diagram of an embodiment of the invention when there is a single TCP source with the parameters mentioned in Fig. 3. The proportions are: the proportion of packets marked red **601**; the proportion of packets marked green **603**; and the proportion of packets marked yellow **605**.

Fig. 7, shows a more complicated simulation – that of packet marking for six TCP flows of a customer using an embodiment of the invention. The proportion of packets marked red for each of the six TCP flows appears in a red group **701**, wherein the proportion is the number of packets marked red as a ratio to the sum of incoming packets of the source. A grouping of six TCP flows marked green **703** are shown each as a proportion of the incoming packets of that TCP flow. A grouping of six TCP flows marked yellow **705** are shown each as a proportion of the incoming packets of that TCP flow.

Fig. 8 is a simulated result of a packet marking for six TCP flows using TCM, a prior art method and apparatus. By comparing Fig. 7 to Fig. 8 it can be seen that compared to TCM, the embodiment of the invention offers better

fairness to flows. A grouping of six TCP flows marked red **801** are shown each as a proportion of the aggregate of the green, yellow and red marked packets of that TCP flow. A grouping of six TCP flows marked green **803** are shown each as a proportion of the aggregate of the green, yellow and red marked packets of that TCP flow. A grouping of six TCP flows marked yellow **805** are shown each as a proportion of the aggregate of the green, yellow and red marked packets of that TCP flow.

Fig. 9 shows a simulation of how a packet burst can be supported by the second embodiment of the invention when the sending rate of a single Constant Bit Rate connection or user is increased from 0.2 Mbps to 1.0 Mbps at time **907**. CIR **911** and PIR **913** are 0.3 Mbps and 0.4 Mbps respectively as before. The throughput of green marked packets **901**, the throughput of yellow marked packets **903**, and the throughput of red marked packets **905**, is shown.

The second embodiment allows the connection or user to use green packets until a burst size is decremented to a level, and then marks the remaining packets based on the CIR and PIR values. Accordingly, a peak green rate occurs **915** that surpasses the CIR and PIR.

Although the invention has been described in the context of particular embodiments, it will be realized that a number of modifications to these teachings may occur to one skilled in the art. A number of metering methods in addition to TSW may be used. Thus, while the invention has been particularly shown and described with respect to specific embodiments thereof, it will be understood by those skilled in the art that changes in form and configuration may be made therein without departing from the scope and spirit of the invention.

CLAIMS

What is claimed is:

1. A method for marking data packets from a source comprising the steps of:

determining a sending rate estimate, s; and

marking a packet to one of a plurality of priority levels based on the sending rate estimate, s.

2. The method of claim 1 wherein the step of marking comprises the steps of:

determining if the sending rate estimate is less than a first rate threshold; and

in response to a determination that the sending rate estimate is less than the first rate threshold, setting a probability of marking at least one data packet with a first selected priority level to a first value, wherein said first selected priority level is one of a plurality of priority levels.

3. The method of claim 2 further comprising the step of:

in response to a determination that the s is less than the first rate threshold, incrementing a burst size.

4. The method of claim 1 wherein the step of marking comprises the steps of:

determining if the sending rate estimate is between a first rate threshold (FRT) and a second rate threshold; and

in response to a determination that the sending rate estimate is between a first rate threshold and a second rate threshold, setting a probability of marking a data packet with a subordinate priority level based on s.

5. The method of claim 1 wherein the step of marking comprises the steps of:

determining if the sending rate estimate is between a first rate threshold (FRT) and a second rate threshold; and

in response to a determination that the sending rate estimate is between a first rate threshold and a second rate threshold, marking a data packet such that a rate of packets marked a subordinate priority level is no greater than 1 - (FRT/s).

6. The method of claim 1 wherein the step of marking comprises the steps of:

determining if the sending rate estimate is above a second rate threshold (SRT); and

in response to a determination that the sending rate estimate is above the SRT, marking the packet such that a rate of packets marked the second priority level is at least (SRT - FRT)/s.

7. The method of claim 6 further comprises the step of:

in response to a determination that the sending rate is above the SRT, marking the packet such that a rate of packets marked a lowest priority level is at least (s-SRT)/s.

8. The method of claim 1 further comprising the steps of:

determining if the sending rate estimate is greater than a rate threshold;

in response to a determination that the sending rate estimate is greater than the rate threshold, determining if a burst size is greater than a minimum burst; and

in response to a determination that the burst size is greater than a minimum burst, marking the packet a first priority level.

9. The method of claim 8 further comprising the step of:

in response to a determination that the burst size is greater than the minimum burst, decrementing the burst size.

10. The method of claim 1 further comprising the steps of:

determining if the sending rate estimate is greater than a super rate threshold;

in response to a determination that the sending rate estimate is greater than the super rate threshold, determining if a burst size is greater than a minimum burst; and

in response to a determination that the burst size is greater than a minimum burst, marking the packet a priority level based on a count of packets marked a highest priority level during a period.

11. The method of claim 10 further comprising the step of:

in response to a determination that the burst size is greater than the minimum burst, decrementing the burst size.

12. An apparatus for marking data packets from a source comprising:

a means for determining a sending rate estimate, s; and

a means for marking a packet to one of a plurality of priority levels based on the sending rate estimate, s.

13. The apparatus of claim 12 wherein the means for marking comprises:

a means for determining if the sending rate estimate is less than a first rate threshold; and

a means for setting a probability of marking at least one data packet with a first selected priority level to a first value, said means responsive to a determination that the sending rate estimate is less than the first rate threshold, wherein said first selected priority level is one of a plurality of priority level.

14. The apparatus of claim 13 further comprises:

a means for incrementing a burst size, in response to a determination that the s is less than the first rate threshold.

15. The apparatus of claim 12 wherein the means for marking comprises:

a means for determining if the sending rate estimate is between a first rate threshold (FRT) and a second rate threshold; and

a means for setting a probability of marking a data packet with a subordinate priority level based on s, said means responsive to a determination that the sending rate estimate is between a first rate threshold and a second rate threshold.

16. The apparatus of claim 12 wherein the means for marking comprises:

a means for determining if the sending rate estimate is between a first rate threshold (FRT) and a second rate threshold; and

a means for marking a data packet such that a rate of packets marked a subordinate priority level is no greater than 1 - (FRT/s) in response to a determination that the sending rate estimate is between a first rate threshold and a second rate threshold.

17. The apparatus of claim 12 wherein the means for marking comprises:

a means for determining if the sending rate estimate is above a second rate threshold (SRT); and

a means for marking the packet such that a rate of packets marked the second priority level is at least (SRT - FRT)/s, in response to a determination that the sending rate estimate is above the SRT.

18. The apparatus of claim 17 further comprises:

a means for marking the packet such that a rate of packets marked a lowest priority level is at least (s-SRT)/s, in response to a determination that the sending rate is above the SRT.

19. The apparatus of claim 12 further comprises:

a means for determining if the sending rate estimate is greater than a rate threshold;

a means for determining if a burst size is greater than a minimum burst, in response to a determination that the sending rate estimate is greater than the rate threshold; and

a means for marking the packet a first priority level, in response to a determination that the burst size is greater than a minimum burst.

20. The apparatus of claim 19 further comprises:

a means for decrementing the burst size, in response to a determination that the burst size is greater than the minimum burst.

21. The apparatus of claim 12 further comprises:

a means for determining if the sending rate estimate is greater than a super rate threshold;

a means for determining if a burst size is greater than a minimum burst, in response to a determination that the sending rate estimate is greater than the super rate threshold; and

a means for marking the packet a priority level based on a count of packets marked a highest priority level during a period, in response to a determination that the burst size is greater than a minimum burst.

22. The apparatus of claim 21 further comprising:

a means for decrementing the burst size, in response to a determination that the burst size is greater than the minimum burst.

23. A method to determine probabilities for marking a packet a priority level comprising the steps of:

determining a first probability;

determining at least one second probability; and

weighting each probability so that each probability contributes to a net probability.

ABSTRACT

A method is disclosed for a router to provide random assignments of three priorities, each signifying a drop precedence, to packets on a per customer basis. The router determines a sending rate estimate. Then the router marks a packet a priority level based on the sending rate estimate.

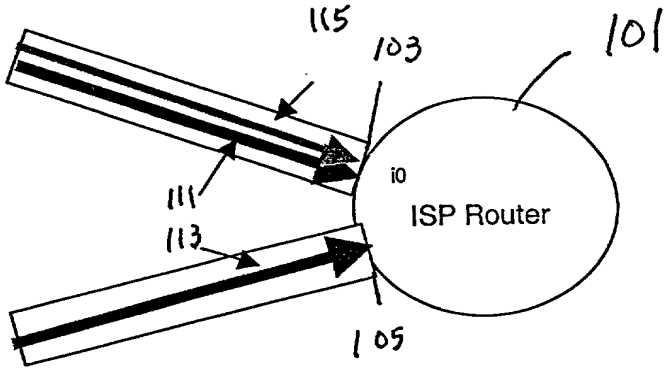


Figure 1

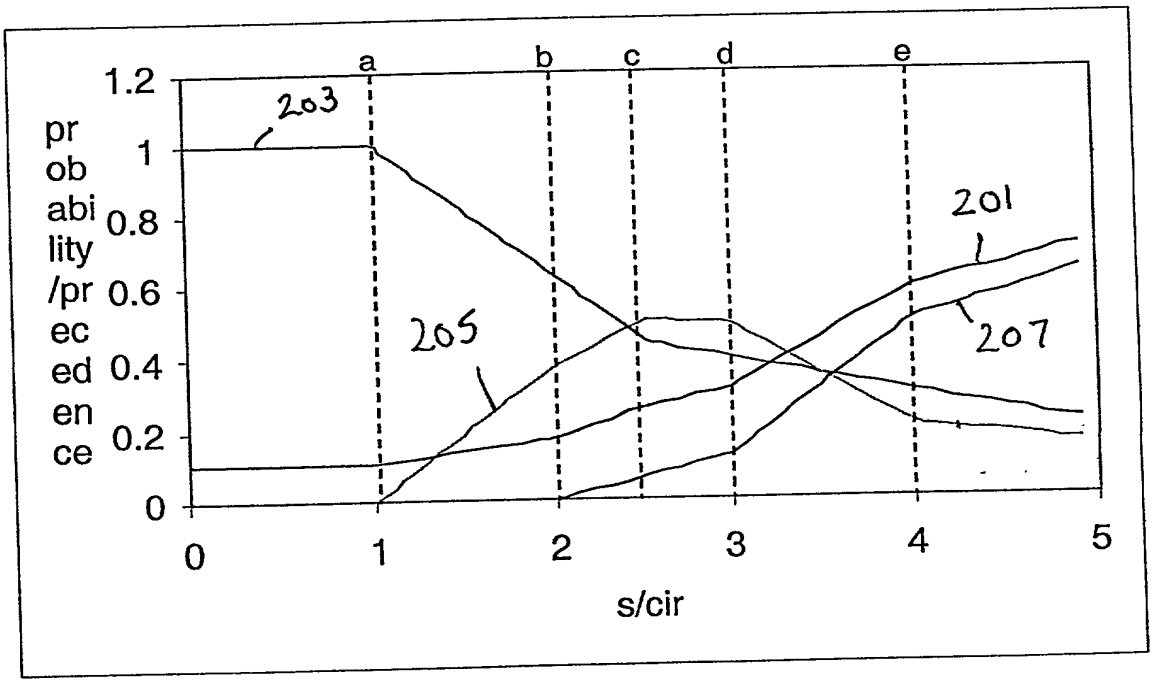


Figure 2

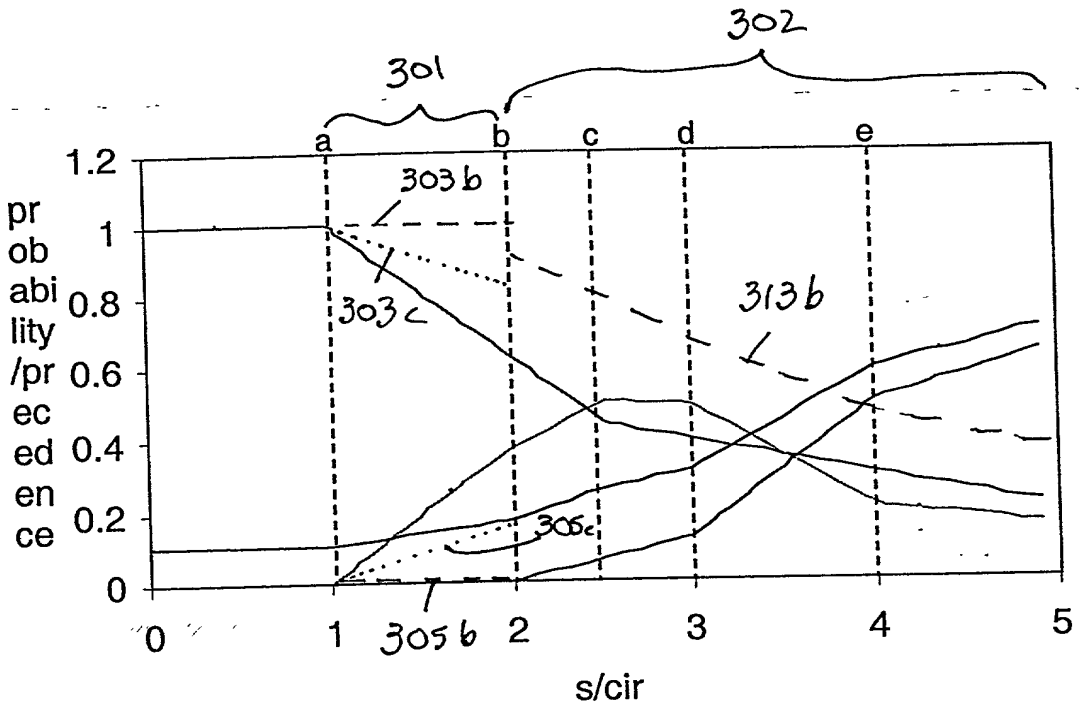

302
301
a
b
c
d
e
1.2
1
0.8
0.6
0.4
0.2
0
probability/preceedence
303b
303c
313b
305c
305b
0
1
2
3
4
5
s/cir

Figure 3

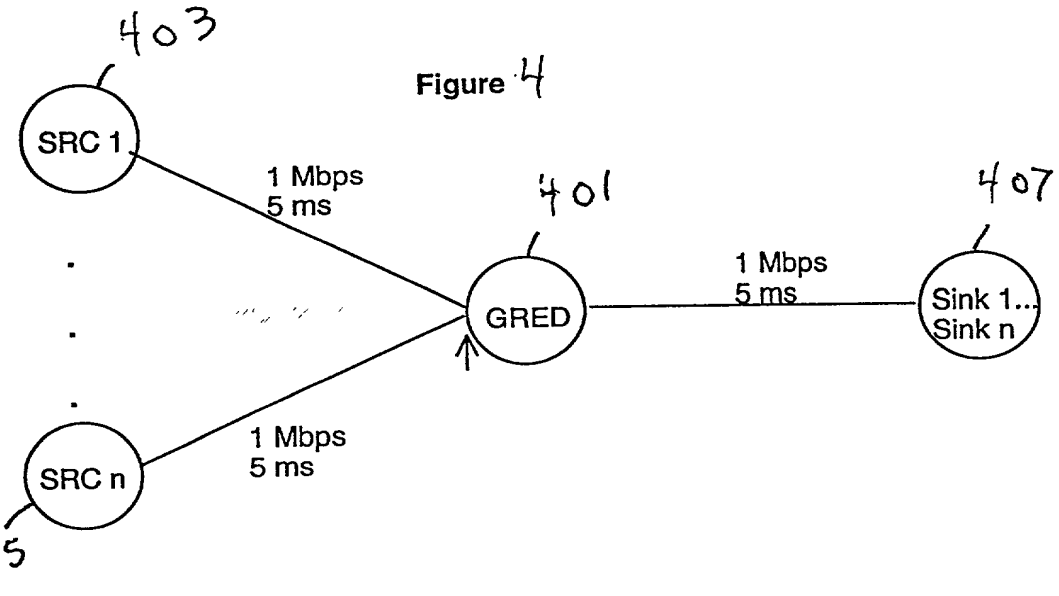


Figure 4

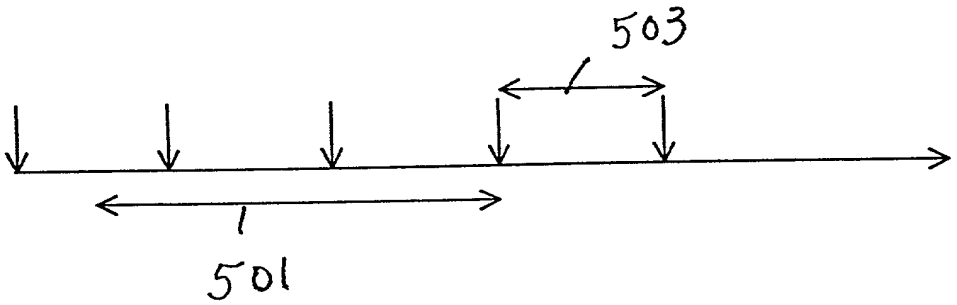


Figure 5

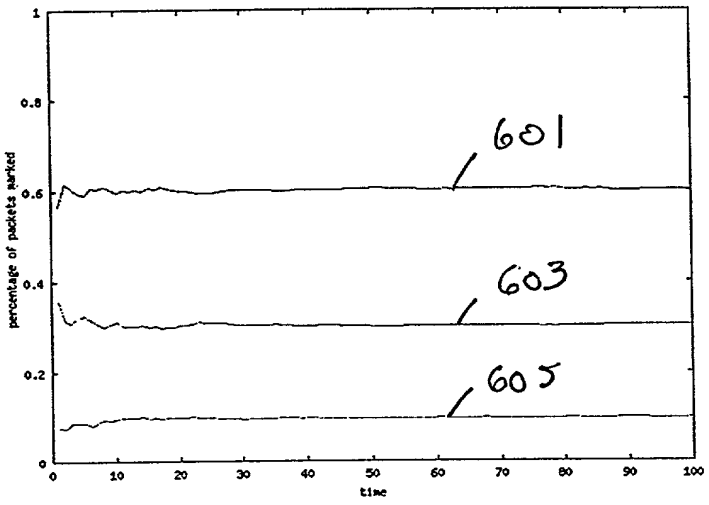


Figure 6

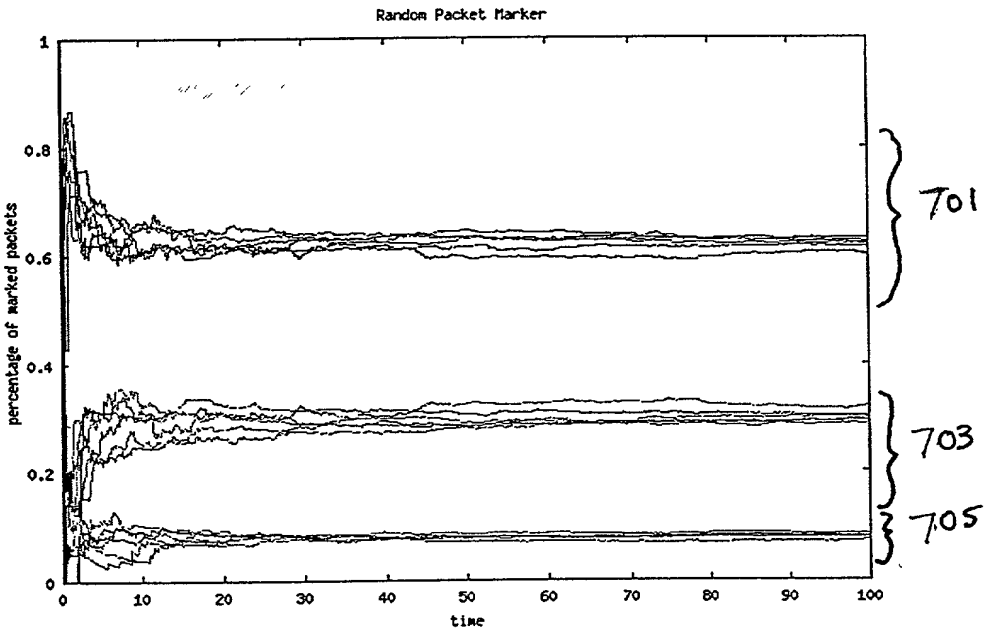

Random Packet Marker
percentage of marked packets
time
0
0.2
0.4
0.6
0.8
1
0
10
20
30
40
50
60
70
80
90
100
701
703
705

Figure 7

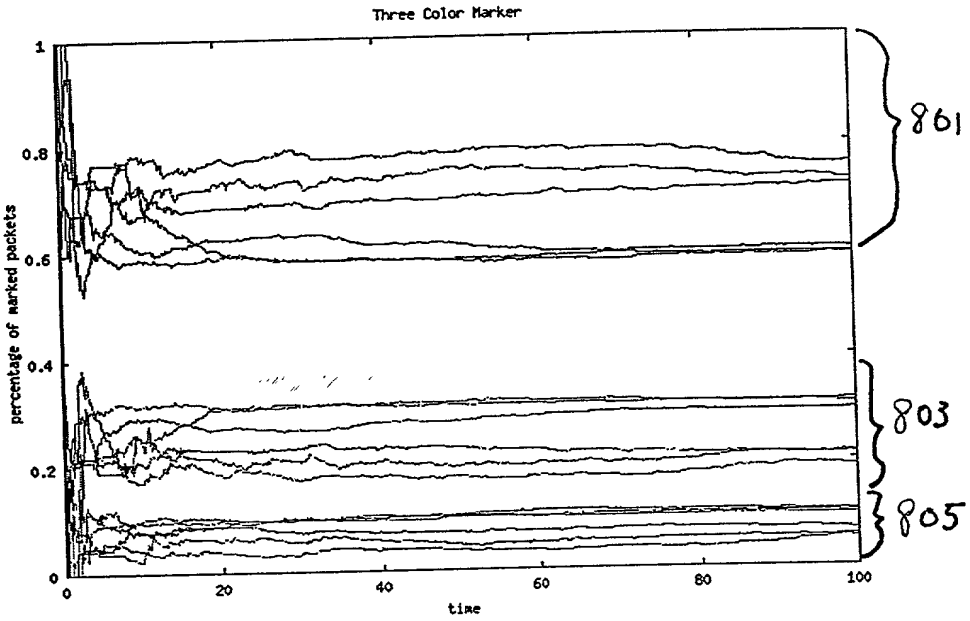

Three Color Marker
percentage of marked packets
time
0
0.2
0.4
0.6
0.8
1
0
20
40
60
80
100
801
803
805

Fig 8

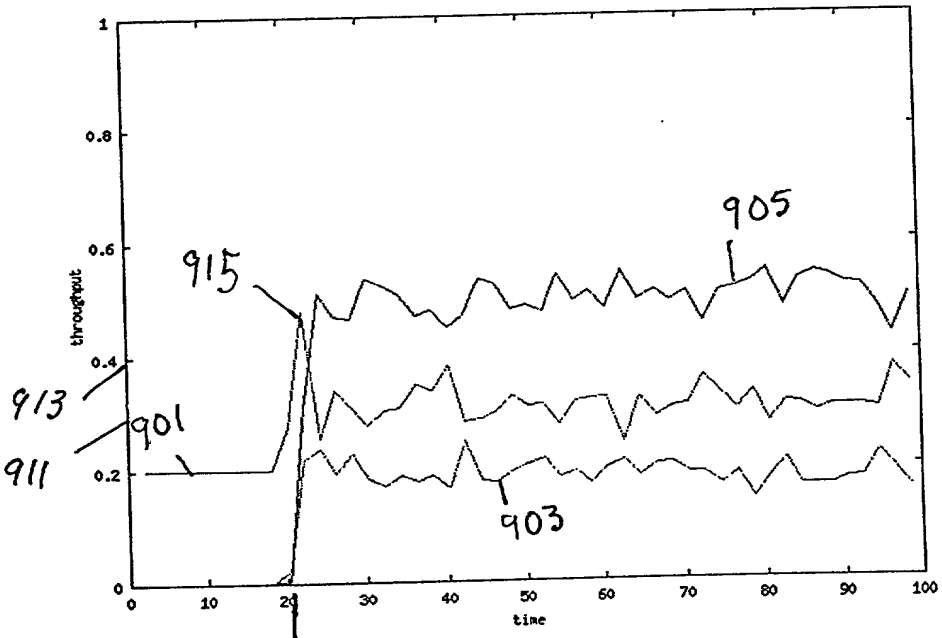

throughput
time
0
0.2
0.4
0.6
0.8
1
0
10
20
30
40
50
60
70
80
90
100
905
915
913
901
911
903
907

Fig. 9